Monge's problem with a quadratic cost by the zero-noise limit of h-path processes

Toshio Mikami*
Department of Mathematics
Hokkaido University
Sapporo 060-0810, Japan
mikami@math.sci.hokudai.ac.jp

Running Head: Zero-noise limit of h-path processes

December 10, 2003

Abstract

We study the asymptotic behavior, in the zero-noise limit, of solutions to Schrödinger's functional equations and that of h-path processes, and give a new proof of the existence of the minimizer of Monge's problem with a quadratic cost.

1 Introduction.

Let $L : \mathbf{R}^d \mapsto [0, \infty)$ be convex, P_0 and P_1 be Borel probability measures on $\mathbf{R}^d$, and put

$$V(P_0, P_1) := \inf\Big\{\int_{\mathbf{R}^d} L(\psi(x) - x)P_0(dx) : P_0\psi^{-1} = P_1\Big\}. \tag{1.1}$$

*Partially supported by the Grant-in-Aid for Scientific Research, No. 15340047 and 15340051, JSPS.

The study of the minimizer of (1.1) can be considered as a special case of Monge's problem.

Kantorovich's approach to Monge's problem is to study the minimizer of the following relaxed problem:

$$\begin{aligned}\tilde{V}(P_0, P_1) \quad := \quad \inf\Big\{ \iint_{\mathbf{R}^d\times\mathbf{R}^d} L(y-x)\mu(dxdy) \qquad\qquad (1.2)\\ : \mu(dx\times\mathbf{R}^d) = P_0(dx), \mu(\mathbf{R}^d\times dy) = P_1(dy)\Big\}.\end{aligned}$$

If there exists a Borel measurable function ψ, on $\mathbf{R}^d$, such that the minimizer of (1.2) is $P_0(dx)\delta_{\psi(x)}(dy)$, then $V(P_0, P_1) = \tilde{V}(P_0, P_1)$ and ψ is a minimizer of (1.1).

This is called the Monge-Kantorovich problem and plays a crucial role in many fields and has been studied by many authors (see [8, 20, 25] and the references therein).

It is easy to see that the following holds:

$$\tilde{V}(P_0, P_1) = \inf\Big\{ E\Big[\int_0^1 L\Big(\frac{d\phi(t)}{dt}\Big)dt\Big]\Big\}, \qquad (1.3)$$

where the infimum is taken over all absolutely continuous stochastic processes $\{\phi(t)\}_{0\le t\le 1}$ for which $P(\phi(t)\in dx) = P_t(dx)$ $(t = 0, 1)$. (In this paper we use the same notation P for different probability measures for the sake of simplicity when it is not confusing.) Indeed, the minimizer of (1.3) is linear in t (see e.g. [5], [10, p. 35]).

This implies that the minimizer of Monge's problem with a quadratic cost $L(u) = |u|^2$ should be the zero-noise limit of h-path processes for Brownian motion, which enables us not to use Kantorovich's approach to study (1.1). By an "h-path process for Brownian motion", we mean an h-path process obtained from a space-time harmonic function of Brownian motion (see (1.7)-(1.10) and also [7, p. 566]).

To make the point clearer, we introduce Schrödinger's functional equation and then briefly describe an h-path process. For $\varepsilon > 0$ and $x \in \mathbf{R}^d$, put

$$g_\varepsilon(x) := \frac{1}{(2\pi\varepsilon)^{d/2}}\exp\Big(-\frac{|x|^2}{2\varepsilon}\Big), \qquad (1.4)$$

$$P_{1,\varepsilon}(dy) := \Big(\int_{\mathbf{R}^d} g_\varepsilon(z-y)P_1(dz)\Big)dy. \tag{1.5}$$

The following is a special case of Schrödinger's functional equations: find nonnegative, σ-finite Borel measures $(\nu_{0,\varepsilon}, \nu_{1,\varepsilon})$ for which

$$\begin{cases} P_0(dx) = \Big(\int_{\mathbf{R}^d} g_\varepsilon(x-y)\nu_{1,\varepsilon}(dy)\Big)\nu_{0,\varepsilon}(dx), \\ P_{1,\varepsilon}(dy) = \Big(\int_{\mathbf{R}^d} g_\varepsilon(x-y)\nu_{0,\varepsilon}(dx)\Big)\nu_{1,\varepsilon}(dy). \end{cases} \tag{1.6}$$

It is known that there exists a unique solution $(\nu_{0,\varepsilon}, \nu_{1,\varepsilon})$ to (1.6) (see [13], and also [22] for the recent development).

Remark 1.1 *(1.6) has a unique solution even if we replace $P_{1,\varepsilon}$ by P_1. But, if we replace $P_{1,\varepsilon}$ by P_1 in (1.6) and consider P_1 as the terminal distribution of the h-path process given below, then we need technical assumptions (A.2) and (A.3) in section 2 (see the discussion below Corollary 2.2 in section 2).*

For $\varepsilon > 0$ and $x \in \mathbf{R}^d$, put

$$h_\varepsilon(t,x) := \begin{cases} \int_{\mathbf{R}^d} g_{\varepsilon(1-t)}(x-y)\nu_{1,\varepsilon}(dy) & (0 \le t < 1), \\ \frac{\nu_{1,\varepsilon}(dx)}{dx} & (t=1). \end{cases} \tag{1.7}$$

Let $(\Omega, \mathbf{B}, P)$ be a probability space, $\{\mathbf{B}_t\}_{t\ge 0}$ be a right continuous, increasing family of sub σ-fields of $\mathbf{B}$, X_o be a $\mathbf{R}^d$-valued, $\mathbf{B}_0$-adapted random variable such that $P(X_o)^{-1} = P_0$, and $\{W(t)\}_{t\ge 0}$ denote a d-dimensional $(\mathbf{B}_t)$-Brownian motion such that $W(0) = o$ (see e.g. [7], [10] or [12]).

The h-path process for $\sqrt{\varepsilon}W(\cdot)$ in $C([0,1])$ with an initial distribution P_0 and a terminal one $P_{1,\varepsilon}$ is the unique weak solution to the following (see [14]): for $t \in [0,1]$,

$$X_\varepsilon(t) = X_o + \int_0^t b_\varepsilon(s, X_\varepsilon(s))ds + \sqrt{\varepsilon}W(t), \tag{1.8}$$

where

$$b_\varepsilon(s,x) := \varepsilon D_x \log h_\varepsilon(s,x) \quad ((s,x) \in [0,1) \times \mathbf{R}^d), \tag{1.9}$$

and $D_x := (\partial/\partial x_i)_{i=1}^d$.

It is known that for any Borel set $A \subset C([0,1])$,

$$P(X_\varepsilon(\cdot) \in A) = E\Big[\frac{h_\varepsilon(1, X_o + \sqrt{\varepsilon}W(1))}{h_\varepsilon(0, X_o)} : X_o + \sqrt{\varepsilon}W(\cdot) \in A\Big]. \qquad (1.10)$$

In particular,

$$P((X_\varepsilon(0), X_\varepsilon(1)) \in dxdy) = \mu_\varepsilon(dxdy) := \nu_{0,\varepsilon}(dx)g_\varepsilon(x-y)\nu_{1,\varepsilon}(dy), \quad (1.11)$$

and $P(X_\varepsilon(1))^{-1} = P_{1,\varepsilon}$.

It is also known that the minimizer of the following is the h-path process in (1.8) (see [11, 26]):

$$V_\varepsilon(P_0, P_{1,\varepsilon}) := \inf\Big\{E\Big[\int_0^1 |u(t)|^2 dt\Big]\Big\}, \qquad (1.12)$$

where the infimum is taken over all $\mathbf{R}^d$-valued, $(\mathbf{B}_t)$-progressively measurable $\{u(t)\}_{0 \le t \le 1}$ for which the distribution of $X_o + \int_0^1 u(s)ds + \sqrt{\varepsilon}W(1)$ is $P_{1,\varepsilon}$, provided that the right hand side of (1.12) is finite.

It seems likely that the h-path process converges, as $\varepsilon \to 0$, to the minimizer of (1.3) with $L(u) = |u|^2$. But it is not trivial that this limit is a function of t and X_o since a continuous strong Markov process which is of bounded variation in time is not always a function of the initial point and time (see [23] and also [19]).

In this paper, independently of known results on the Monge-Kantorovich problem, we show that $V_\varepsilon(P_0, P_{1,\varepsilon})$ converges to $V(P_0, P_1)$ and $X_\varepsilon(1)$ converges, in L^2, to the minimizer of (1.1) as $\varepsilon \to 0$, when $L(u) = |u|^2$. As a by-product, we give a new proof of the existence of the minimizer of (1.1) with $L(u) = |u|^2$.

From a probabilistic interest, replacing $P_{1,\varepsilon}$ by P_1 in (1.6)-(1.12), we also show the similar result to above, under technical assumptions.

If $P_0(dx)$ is absolutely continuous with respect to dx (see (A.1) in section 2) and $L(u) = |u|^2$, then it is known that (1.1) and (1.2) have the unique minimizers $D\varphi(x)$ and $P_0(dx)\delta_{D\varphi(x)}(dy)$ respectively, where $\varphi : \mathbf{R}^d \mapsto (-\infty, \infty]$ is convex (see [3, 4], and also [8, 15, 16, 20, 21, 25] and the reference therein, and also [18, 19] for the continuum limit of (1.3)).

When $L(u) = |u|$, in [9] they studied (1.2) by the "$p \to \infty$" limit of the minimization problem for which the Euler-Lagrange equation is the p-Laplacian PDE under the assumption that P_0 and P_1 have disjoint compact supports, and in [6] and [24] they studied (1.2) by the "$q \downarrow 1$" limit of (1.2) with $L(u) = |u|^q$ under the assumption that P_0 and P_1 have compact supports (see also [1]).

In future we would like to study the zero noise limit of the minimizer of (1.12) with a more general cost function $L(u)$, instead of $|u|^2$, and then apply the result to Monge's problem.

In section 2 we give our main result which will be proved in section 3.

2 Main Result.

In this section we give our main result. We first state assumptions.
(A.0) P_0 and P_1 are Borel probability measures, on $\mathbf{R}^d$, which have finite second moments, i.e.,

$$\int_{\mathbf{R}^d} |x|^2 (P_0(dx) + P_1(dx)) < \infty.$$

(A.1) $p_0(x) := P_0(dx)/dx$ exists.

Then the following holds.

Theorem 2.1 *Suppose that (A.0) holds. Then $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ is tight, and any weak limit point of $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ as $\varepsilon \to 0$ is supported on a cyclically monotone set.*

For the readers' convenience, we introduce the following.

Definition 2.1 *The nonempty set $A \in \mathbf{R}^d \times \mathbf{R}^d$ is called cyclically monotone if for any $n \geq 1$ and any $(x_i, y_i) \in A$ $(i = 1, \cdots, n)$,*

$$\sum_{i=1}^{n} < y_i, x_{i+1} - x_i > \leq 0 \tag{2.1}$$

(see e.g. [25, p. 80]), where $x_{n+1} := x_1$, and $< \cdot, \cdot >$ denotes the inner product in $\mathbf{R}^d$.

Since a cyclically monotone set in $\mathbf{R}^d \times \mathbf{R}^d$ is contained in the subdifferential of a proper lower semicontinuous convex function on $\mathbf{R}^d$ and since a proper convex function is differentiable dx-a.e. in the interior of its domain (see [25, pp. 52, 82]), we obtain the following.

Corollary 2.1 *Suppose that (A.0) and (A.1) hold. Then for any weak limit point μ of $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ as $\varepsilon \to 0$, there exists a proper lower semicontinuous convex function $\varphi : \mathbf{R}^d \mapsto (-\infty, \infty]$ such that*

$$\mu(dxdy) = P_0(dx)\delta_{D\varphi(x)}(dy). \tag{2.2}$$

Remark 2.1 *If (A.1) holds and $p_1(y) := P_1(dy)/dy$ exists (see (A.2) given later), then Corollary 2.1 gives a new proof of the existence to the following Monge-Ampère equation:*

$$p_0(x) = p_1(D\varphi(x)) \det(D^2\varphi(x)) \tag{2.3}$$

in the sense that $P_0(D\varphi)^{-1} = P_1$, where $D^2 := (\partial^2/\partial x_i \partial x_j)_{i,j=1}^d$. For the regularity results on the solution to (2.3), see [25, pp. 140-141], [5, Theorem 1.1] and the references therein.

The following which can be proved from Theorem 2.1 and Corollary 2.1, independently of known results on the Monge-Kantorovich problem [1, 3, 4, 15, 16, 21], is our main result.

Theorem 2.2 *Suppose that (A.0) and (A.1) hold, and that $L(u) = |u|^2$. Then*

$$\lim_{\varepsilon\to 0} V_\varepsilon(P_0, P_{1,\varepsilon}) = V(P_0, P_1) < \infty. \tag{2.4}$$

In particular, $D\varphi$ in Corollary 2.1 is the unique minimizer of (1.1), and the following holds:

$$\lim_{\varepsilon\to 0} E[\int_0^1 |b_\varepsilon(t, X_\varepsilon(t)) - (D\varphi(X_o) - X_o)|^2 dt] = 0, \tag{2.5}$$

$$\lim_{\varepsilon\to 0} E[\sup_{0\le t\le 1} |X_\varepsilon(t) - \{X_o + t(D\varphi(X_o) - X_o)\}|^2] = 0. \tag{2.6}$$

The following is known on (1.1)-(1.3) with $L(u) = |u|^2$.

(i) Suppose that (A.0) holds. Then a probability measure supported on a cyclically monotone set in $\mathbf{R}^d \times \mathbf{R}^d$ is a minimizer of (1.2) (see [15, 16] and also [25, pp. 66, 82], [1, Theorem 3.2]).

(ii) Suppose that (A.0) and (A.1) hold. Then there exists a convex function φ such that $P_0(dx)\delta_{D\varphi(x)}(dy)$ is the unique minimizer of (1.2) (see [3, 4]).

Using these facts, we have the following.

Corollary 2.2 *(i) Suppose that (A.0) holds and that $L(u) = |u|^2$. Then any weak limit point of $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ as $\varepsilon \to 0$ is a minimizer of (1.2). (ii) Suppose in addition that (A.1) holds. Then μ_ε weakly converges to the unique minimizer of (1.2) as $\varepsilon \to 0$.*

In (1.6) we considered $P_{1,\varepsilon}$, instead of P_1, to avoid technical assumptions in Theorem 2.2 (see Remark 1.1). As far as the zero-noise limit of h-path processes is concerned, there is no reason to perturb the terminal distribution P_1. From a probabilistic interest, we discuss the zero-noise limit of h-path processes for Brownian motion with the terminal distribution P_1.

Repalce $P_{1,\varepsilon}$ by P_1 in (1.6). Then there exists a unique solution $(\overline{\nu}_{0,\varepsilon}, \overline{\nu}_{1,\varepsilon})$ to (1.6) (see [13]). We define $\overline{\mu}_\varepsilon$ from $(\overline{\nu}_{0,\varepsilon}, \overline{\nu}_{1,\varepsilon})$ in the same way as in (1.11).

If we assume

(A.2) $p_1(x) := P_1(dx)/dx$ exists,

then we can define $\overline{h}_\varepsilon(t,x)$, $\overline{X}_\varepsilon(t)$ and $\overline{b}_\varepsilon(s,x)$ in the same way as in (1.7)-(1.9), respectively, by replacing $(\nu_{0,\varepsilon}, \nu_{1,\varepsilon})$ by $(\overline{\nu}_{0,\varepsilon}, \overline{\nu}_{1,\varepsilon})$ (see [14]).

If we assume in addition that the following holds:

(A.3) $\int_{\mathbf{R}^d} (\log p_1(x)) P_1(dx) < \infty$,

then $V_\varepsilon(P_0, P_1)$ is finite for $\varepsilon > 0$ and the similar result to Theorem 2.2 and Corollary 2.2 holds for $\overline{X}_\varepsilon$. More precisely, the following holds.

Proposition 2.1 *(i) Suppose that (A.0) holds and that $L(u) = |u|^2$. Then $\{\overline{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ is tight and any weak limit point of $\{\overline{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ as $\varepsilon \to 0$ is a minimizer of (1.2). (ii) Suppose in addition that (A.1) holds. Then $\overline{\mu}_\varepsilon$ weakly converges to the unique minimizer of (1.2) as $\varepsilon \to 0$. (iii) Suppose in addition that (A.2) and (A.3) hold. Then $V_\varepsilon(P_0, P_1)$ is finite for $\varepsilon > 0$, and*

$$\lim_{\varepsilon\to 0} V_\varepsilon(P_0, P_1) = V(P_0, P_1), \tag{2.7}$$

and for $D\varphi$ in Corollary 2.1,

$$\lim_{\varepsilon\to 0} E[\sup_{0\le t\le 1} |\overline{X}_\varepsilon(t) - \{X_o + t(D\varphi(X_o) - X_o)\}|^2] = 0. \tag{2.8}$$

3 Proof.

In this section we prove our results stated in section 2.

We first state and prove technical lemmas to prove Theorem 2.1. For x, $y \in \mathbf{R}^d$, $m \geq 1$ and $\varepsilon > 0$, put

$$\begin{aligned} H_{m,\varepsilon}(x,y) \quad &:= \quad \varepsilon \log\Big\{ \iint_{U_m(o)\times U_m(o)} \exp\Big(\frac{< x, z_1 > + < y, z_0 >}{\varepsilon} \\ &\qquad - \frac{< z_0, z_1 >}{\varepsilon} \Big) \mu_\varepsilon(dz_0 dz_1) \Big\}, \end{aligned} \tag{3.1}$$

$$H_{i,m,\varepsilon}(x) := \varepsilon \log\Big(\int_{U_m(o)} g_\varepsilon(x - z_j)\nu_{j,\varepsilon}(dz_j) \Big) + \frac{|x|^2}{2} \ (i, j = 0, 1, \ i \neq j), \tag{3.2}$$

$$\mu_{0,m,\varepsilon}(dz_0) := \mu_\varepsilon(dz_0 \times U_m(o)), \quad \mu_{1,m,\varepsilon}(dz_1) := \mu_\varepsilon(U_m(o) \times dz_1) \tag{3.3}$$

(see (1.11)), where $U_m(o) := \{x \in \mathbf{R}^d : |x| < m\}$. Then the following holds.

Lemma 3.1 *(i) For any $m \geq 1$ and $\varepsilon > 0$ for which $\mu_\varepsilon(U_m(o) \times U_m(o)) > 0$, and any x and $y \in \mathbf{R}^d$,*

$$H_{m,\varepsilon}(x,y) = H_{0,m,\varepsilon}(x) + H_{1,m,\varepsilon}(y) + \varepsilon \log(2\pi\varepsilon)^{d/2}, \tag{3.4}$$

$$\mu_\varepsilon(dz_0 dz_1) = \exp\Big(\frac{1}{\varepsilon}(< z_0, z_1 > - H_{m,\varepsilon}(z_0, z_1)) \Big) \mu_{0,m,\varepsilon}(dz_0)\mu_{1,m,\varepsilon}(dz_1), \tag{3.5}$$

$$\begin{aligned} H_{m,\varepsilon}(x,y) &= \varepsilon \log\Big\{ \iint_{U_m(o)\times U_m(o)} \exp\Big(\frac{< x, z_1 > + < y, z_0 >}{\varepsilon} \\ &\qquad - \frac{H_{m,\varepsilon}(z_0, z_1)}{\varepsilon} \Big) \mu_{0,m,\varepsilon}(dz_0)\mu_{1,m,\varepsilon}(dz_1) \}. \end{aligned} \tag{3.6}$$

(ii) For any $m \geq 1$ and $\varepsilon > 0$ for which $\mu_\varepsilon(U_m(o) \times U_m(o)) > 0$, $H_{m,\varepsilon}(\cdot, \cdot)$ is convex, and for any x and $y \in \mathbf{R}^d$,

$$|H_{m,\varepsilon}(x,y)| \leq (|x| + |y|)m + m^2 - \varepsilon \log \mu_\varepsilon(U_m(o) \times U_m(o)). \tag{3.7}$$

Proof. We first prove (i). (3.4) can be obtained from (1.11) and (3.1)-(3.2) easily. (3.5) holds from (1.11), (3.4) and from the following: for i, $j = 0$, 1 for which $i \neq j$,

$$\frac{\mu_{i,m,\varepsilon}(dz_i)}{\nu_{i,\varepsilon}(dz_i)} = \int_{U_m(o)} g_\varepsilon(z_i - z_j)\nu_{j,\varepsilon}(dz_j) = \exp\Big(\frac{1}{\varepsilon}\Big(H_{i,m,\varepsilon}(z_i) - \frac{|z_i|^2}{2}\Big)\Big). \quad (3.8)$$

(3.6) can be obtained from (3.1) and (3.5) easily.

Next we prove (ii). $H_{m,\varepsilon}(\cdot,\cdot)$ is convex since for any $\lambda \in (0,1)$ and any (x,y), $(\tilde{x},\tilde{y}) \in \mathbf{R}^d \times \mathbf{R}^d$,

$$\begin{aligned}
& H_{m,\varepsilon}(\lambda x + (1-\lambda)\tilde{x}, \lambda y + (1-\lambda)\tilde{y}) \\
= \ & \varepsilon \log\Big\{\iint_{U_m(o)\times U_m(o)} \exp\Big(\frac{\lambda(<x,z_1> + <y,z_0> - <z_0,z_1>)}{\varepsilon}\Big) \\
& \quad \times \exp\Big(\frac{(1-\lambda)(<\tilde{x},z_1> + <\tilde{y},z_0> - <z_0,z_1>)}{\varepsilon}\Big)\mu_\varepsilon(dz_0 dz_1)\Big\} \\
\le \ & \lambda H_{m,\varepsilon}(x,y) + (1-\lambda)H_{m,\varepsilon}(\tilde{x},\tilde{y})
\end{aligned}$$

by Hölder's inequality. (3.7) can be obtained from (3.1) easily.

Q. E. D.

Remark 3.1 *For $x \in \mathbf{R}^d$, $m \ge 1$, $\varepsilon > 0$, and i, $j = 0$, 1 $(i \ne j)$,*

$$H_{i,m,\varepsilon}(x) = \varepsilon \log\Big(\int_{U_m(o)} \frac{1}{(2\pi\varepsilon)^{d/2}} \exp\Big(\frac{1}{\varepsilon}(<x,z_j> - H_{j,m,\varepsilon}(z_j))\Big)\mu_{j,m,\varepsilon}(dz_j)\Big)$$

from (3.2) and (3.8).

Lemma 3.2 *Suppose that (A.0) holds. Then for any sequence $\{\varepsilon_n\}_{n\ge1}$ for which $\varepsilon_n \to 0$ as $n \to \infty$, there exist a subsequence $\{\varepsilon_{n(k)}\}_{k\ge1}$ and $m_0 \ge 1$ such that $H_{m,\varepsilon_{n(k)}}$ is convergent in $C(\mathbf{R}^d \times \mathbf{R}^d)$ as $k \to \infty$ for all $m \ge m_0$. In particular,*

$$m \mapsto H_m := \lim_{k\to\infty} H_{m,\varepsilon_{n(k)}}(: \mathbf{R}^d \times \mathbf{R}^d \mapsto (-\infty,\infty)) \quad (3.9)$$

is nondecreasing on $\{m_0, m_0+1, \cdots\}$,

$$(x,y) \mapsto H(x,y) := \lim_{m\to\infty} H_m(x,y) \in (-\infty,\infty] \quad (3.10)$$

is convex on $\mathbf{R}^d \times \mathbf{R}^d$,

$$< x, y > -H(x, y) \leq 0 \quad ((x, y) \in supp(P_0) \times supp(P_1)), \tag{3.11}$$

and the following set is cyclically monotone:

$$S := \{(x, y) \in supp(P_0) \times supp(P_1) | < x, y >= H(x, y)\}. \tag{3.12}$$

Proof. There exist $m_0 \geq 1$ such that for any $m \geq m_0$, $\{H_{m,\varepsilon_n}\}_{n\geq 1}$ is bounded in $U_{\ell+1}(o) \times U_{\ell+1}(o)$ for any $\ell \geq 1$ from (3.7) and from the following:

$$\begin{aligned}
& 1 - \mu_\varepsilon(U_m(o) \times U_m(o)) \qquad\qquad (3.13)\\
\leq \; & \frac{\iint_{\mathbf{R}^d \times \mathbf{R}^d}(|x|^2 + |y|^2)\mu_\varepsilon(dxdy)}{m^2} = \frac{\int_{\mathbf{R}^d} |x|^2 P_0(dx) + \int_{\mathbf{R}^d} |y|^2 P_{1,\varepsilon}(dy)}{m^2} \\
= \; & \frac{\int_{\mathbf{R}^d} |x|^2 P_0(dx) + \iint_{\mathbf{R}^d \times \mathbf{R}^d} |x + y|^2 g_\varepsilon(x) dx P_1(dy)}{m^2} \\
\leq \; & \frac{\int_{\mathbf{R}^d} |x|^2 P_0(dx) + 2(\varepsilon d + \int_{\mathbf{R}^d} |y|^2 P_1(dy))}{m^2} \to 0 \quad (\text{as } m \to \infty)
\end{aligned}$$

from (A.0), uniformly for $\varepsilon = \varepsilon_n$ $(n \geq 1)$. Hence for any $m \geq m_0$ and any $\ell \geq 1$, $\{H_{m,\varepsilon_n}\}_{n\geq 1}$ contains a uniformly convergent subsequence on $U_\ell(o) \times U_\ell(o)$ since $H_{m,\varepsilon_n}(\cdot, \cdot)$ is convex from Lemma 3.1, (ii) (see [2, p. 21, Theorem 3.2]). By the diagonal method, $\{H_{m,\varepsilon_n}\}_{n\geq 1}$ contains a convergent subsequence $\{H_{m,\varepsilon_{m,n}}\}_{n\geq 1}$ in $C(\mathbf{R}^d \times \mathbf{R}^d)$. In particular, we can take $\{\varepsilon_{m,n}\}_{n\geq 1}$ so that $m \mapsto \{\varepsilon_{m,n}\}_{n\geq 1}$ is nonincreasing on $\{m_0, m_0 + 1, \cdots\}$. Put

$$\varepsilon_{n(k)} := \varepsilon_{k+m_0-1,k+m_0-1} \quad (k \geq 1).$$

Then $H_{m,\varepsilon_{n(k)}}$ is convergent in $C(\mathbf{R}^d \times \mathbf{R}^d)$ as $k \to \infty$ for all $m \geq m_0$.

$m \mapsto H_m$ is nondecreasing on $\{m_0, m_0 + 1, \cdots\}$ since

$$H_{m+1,\varepsilon_{n(k)}} \geq H_{m,\varepsilon_{n(k)}} \quad (k \geq 1)$$

for all $m \geq m_0$ from (3.1). Hence for any $(x, y) \in \mathbf{R}^d \times \mathbf{R}^d$, $H_m(x, y)$ is convergent or diverges to ∞ as $m \to \infty$.

As the limit of convex functions, $H(\cdot, \cdot)$ in (3.10) is convex in $\mathbf{R}^d \times \mathbf{R}^d$.

For any $(x,y)\in$ supp$(P_0)\times$supp(P_1), $r>0$, $m\ge r+|x|+|y|+m_0$ and $k\ge 1$, from (3.6),

$$\begin{aligned}
& H_{m,\varepsilon_{n(k)}}(x,y) \qquad (3.14)\\
\ge\ & \inf_{(z_0,z_1)\in U_r(x)\times U_r(y)}\{<x,z_1>+<y,z_0>-H_{m,\varepsilon_{n(k)}}(z_0,z_1)\}\\
& +\varepsilon_{n(k)}\log\{\mu_{0,m,\varepsilon_{n(k)}}(U_r(x))\mu_{1,m,\varepsilon_{n(k)}}(U_r(y))\}.
\end{aligned}$$

Since $H_{m,\varepsilon_{n(k)}}$ converges to H_m as $k\to\infty$, uniformly on every compact subset of $\mathbf{R}^d\times\mathbf{R}^d$,

$$\begin{aligned}
& \inf_{(z_0,z_1)\in U_r(x)\times U_r(y)}(<x,z_1>+<y,z_0>-H_{m,\varepsilon_{n(k)}}(z_0,z_1)) \qquad (3.15)\\
\to\ & \inf_{(z_0,z_1)\in U_r(x)\times U_r(y)}(<x,z_1>+<y,z_0>-H_m(z_0,z_1)) \quad (\text{as } k\to\infty)\\
\to\ & 2<x,y>-H_m(x,y) \quad (\text{as } r\to 0)\\
\to\ & 2<x,y>-H(x,y) \quad (\text{as } m\to\infty).
\end{aligned}$$

From (A.0), for sufficiently large $m\ge 1$,

$$\liminf_{\varepsilon\to 0}\{\mu_{0,m,\varepsilon}(U_r(x))\mu_{1,m,\varepsilon}(U_r(y))\}>0. \qquad (3.16)$$

Indeed,

$$\begin{aligned}
& \mu_{0,m,\varepsilon}(U_r(x))\mu_{1,m,\varepsilon}(U_r(y))\\
=\ & \{P_0(U_r(x))-\mu_\varepsilon(U_r(x)\times U_m(o)^c)\}\{P_{1,\varepsilon}(U_r(y))-\mu_\varepsilon(U_m(o)^c\times U_r(y))\}.
\end{aligned}$$

$$\mu_\varepsilon(U_r(x)\times U_m(o)^c)\le\frac{1}{m^2}\int_{\mathbf{R}^d}|z|^2P_{1,\varepsilon}(dz)\le\frac{2(\varepsilon d+\int_{\mathbf{R}^d}|z|^2P_1(dz))}{m^2}$$

as in (3.13), and

$$\mu_\varepsilon(U_m(o)^c\times U_r(y))\le\frac{1}{m^2}\int_{\mathbf{R}^d}|z|^2P_0(dz),$$

$$\liminf_{\varepsilon\to 0}P_{1,\varepsilon}(U_r(y))\ge P_1(U_r(y))$$

since $P_{1,\varepsilon}$ weakly converges to P_1 as $\varepsilon \to 0$. Hence (3.16) holds since for $(x,y) \in \operatorname{supp}(P_0) \times \operatorname{supp}(P_1)$,

$$(P_0 \times P_1)(U_r(x) \times U_r(y)) > 0.$$

(3.14)-(3.16) implies (3.11).

The set S is cyclically monotone. Indeed, for any $k, \ell \geq 1$, $(x_1, y_1), \cdots, (x_\ell, y_\ell) \in S$ and $m \geq m_0$, putting $x_{\ell+1} := x_1$,

$$\sum_{i=1}^{\ell} (H_{m,\varepsilon_{n(k)}}(x_{i+1}, y_i) - H_{m,\varepsilon_{n(k)}}(x_i, y_i)) = 0 \tag{3.17}$$

from (3.4). Let $k \to \infty$ and then $m \to \infty$. Then from (3.11),

$$\sum_{i=1}^{\ell} < y_i, x_{i+1} - x_i > \leq \sum_{i=1}^{\ell} (H(x_{i+1}, y_i) - H(x_i, y_i)) = 0. \tag{3.18}$$

(Notice that $H(x_i, y_i)$ is finite for all $i = 1, \cdots, \ell$.)

Q. E. D.

Remark 3.2 *H is lower semicontinuous since $H_m \uparrow H$ as $m \to \infty$ and since $H_m \in C(\mathbf{R}^d \times \mathbf{R}^d)$ as a finite convex function for sufficiently large m (see the proof of Lemma 3.2). If $H(x,y)$ and $H(a,b)$ are finite, then $H(x,b)$ and $H(a,y)$ are also finite since for sufficienlty large $m \geq 1$, from (3.10) and (3.17),*

$$-\infty < H_m(x,b) + H_m(a,y) \leq H(x,b) + H(a,y) = H(x,y) + H(a,b) < \infty.$$

In particular,

$$H(x,y) = H(x,b) + H(a,y) - H(a,b).$$

(Proof of Theorem 2.1.) $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ is tight from (3.13) (see e.g. [12, p. 7]). Take a weakly convergent subsequence $\{\mu_{\varepsilon_n}\}_{n \geq 1}$ and denote by μ its weak limit, where $\varepsilon_n \to 0$ as $n \to \infty$.

By taking $m_0 \geq 1$ and a subsequence $\{\varepsilon_{n(k)}\}_{k \geq 1}$, construct a convex function H as in Lemma 3.2.

From (3.11)-(3.12), we only have to show the following to complete the proof:

$$\mu(\{(x,y)| < x,y > -H(x,y) < 0\}) = 0. \tag{3.19}$$

By the monotone convergence theorem and Lemma 3.2,

$$\begin{aligned} & \mu(\{(x,y)| < x,y > -H(x,y) < 0\}) \\ = & \lim_{r\downarrow 0}(\lim_{m\uparrow\infty} \mu(\{(x,y)| < x,y > -H_m(x,y) < -r\})). \end{aligned} \tag{3.20}$$

For any $m \geq m_0$, $H_{m,\varepsilon_{n(k)}}$ converges to H_m as $k \to \infty$, uniformly on every compact subset of $\mathbf{R}^d \times \mathbf{R}^d$. Therefore for any $R > 0$,

$$\begin{aligned} & \mu(\{(x,y)| < x,y > -H_m(x,y) < -r, |x|, |y| < R\}) \\ \leq & \liminf_{k\to\infty} \mu_{\varepsilon_{n(k)}}(\{(x,y)| < x,y > -H_m(x,y) < -r, |x|, |y| < R\}) \\ \leq & \liminf_{k\to\infty} \mu_{\varepsilon_{n(k)}}(\{(x,y)| < x,y > -H_{m,\varepsilon_{n(k)}}(x,y) < -r/2, |x|, |y| < R\}) \\ \leq & \liminf_{k\to\infty} \exp\Big(-\frac{r}{2\varepsilon_{n(k)}}\Big) = 0 \quad \text{(from (3.5))}. \end{aligned} \tag{3.21}$$

Notice that the set $\{(x,y)| < x,y > -H_m(x,y) < -r, |x|, |y| < R\}$ is open since $H_m \in C(\mathbf{R}^d \times \mathbf{R}^d)$ from Lemma 3.1, (ii).

Letting $R \to \infty$ in (3.21), we obtain (3.19) from (3.20).

Q. E. D.

Next we prove Theorem 2.2.

(Proof of Theorem 2.2). The proof of (2.4) is devided into the following:

$$\liminf_{\varepsilon\to 0} V_\varepsilon(P_0, P_{1,\varepsilon}) \geq V(P_0, P_1), \tag{3.22}$$

$$\limsup_{\varepsilon\to 0} V_\varepsilon(P_0, P_{1,\varepsilon}) \leq V(P_0, P_1) < \infty. \tag{3.23}$$

To prove (3.22), we only have to show that for any $\{\varepsilon_n\}_{n\geq 1}$ for which $\varepsilon_n \to 0$ and $E[\int_0^1 |b_{\varepsilon_n}(s, X_{\varepsilon_n}(s))|^2 ds]$ is convergent as $n \to \infty$,

$$\lim_{n\to\infty} E[\int_0^1 |b_{\varepsilon_n}(s, X_{\varepsilon_n}(s))|^2 ds] \geq V(P_0, P_1) \tag{3.24}$$

(see (1.8) for notation). (3.24) holds since $\{X_{\varepsilon_n}(\cdot)\}_{n\geq 1}$ is tight in $C([0,1])$, and since any weak limit point $X(\cdot)$ of $\{X_{\varepsilon_n}(\cdot)\}_{n\geq 1}$ is an absolutely continuous stochastic process (see e.g. [19, Lemmas 2-3]), and

$$\begin{aligned}&\lim_{n\to\infty} E[\int_0^1 |b_{\varepsilon_n}(s, X_{\varepsilon_n}(s))|^2 ds] \\ \geq\ & E[\int_0^1 \left|\frac{dX(s)}{ds}\right|^2 ds] \geq E[|X(1)-X(0)|^2] \geq V(P_0, P_1)\end{aligned} \tag{3.25}$$

from (1.11) and (2.2) (see e.g. [19, the proof of (3.17)]).

Next we prove (3.23). Take ψ for which $P_0\psi^{-1} = P_1$, which is possible from Corollary 2.1. Then from (A.0),

$$V(P_0, P_1) \leq E[|\psi(X_o) - X_o|^2] \leq 2\int_{\mathbf{R}^d} |x|^2 (P_0(dx) + P_1(dx)) < \infty. \tag{3.26}$$

Put

$$X_{\varepsilon,\psi}(t) := X_o + t(\psi(X_o) - X_o) + \sqrt{\varepsilon} W(t). \tag{3.27}$$

Then $P(X_{\varepsilon,\psi}(1))^{-1} = P_{1,\varepsilon}$, which implies (3.23).

By (2.2), (2.4) and (3.25), $D\varphi$ in Corollary 2.1 is a minimizer of (1.1) with $L(u) = |u|^2$. The uniqueness of the minimizer of (1.1) with $L(u) = |u|^2$ is can be shown easily (see e.g. [25, p. 69]).

(2.5)-(2.6) is an easy consequence of (2.4). For $t \in [0,1]$,

$$\begin{aligned}&|X_\varepsilon(t) - \{X_o + t(D\varphi(X_o) - X_o)\}| \\ \leq\ & \int_0^1 |b_\varepsilon(s, X_\varepsilon(s)) - (D\varphi(X_o) - X_o)| ds + \sqrt{\varepsilon} \sup_{0\leq t\leq 1} |W(t)|.\end{aligned} \tag{3.28}$$

$$E[\sup_{0\leq t\leq 1} |W(t)|^2] \leq 4d \tag{3.29}$$

(see e.g. [12, p. 34]), and from (2.4),

$$E[\int_0^1 |b_\varepsilon(s, X_\varepsilon(s)) - (D\varphi(X_o) - X_o)|^2 ds] \tag{3.30}$$

$$
\begin{aligned}
&= E[\int_0^1 |b_\varepsilon(s,X_\varepsilon(s))|^2 ds + |D\varphi(X_o)-X_o|^2] \\
&\qquad -2E[< X_\varepsilon(1) - X_o - \sqrt{\varepsilon}W(1), D\varphi(X_o)-X_o >] \\
&\to 2V(P_0,P_1) - 2E[< D\varphi(X_o)-X_o, D\varphi(X_o)-X_o >] = 0 \quad \text{as } \varepsilon \to 0.
\end{aligned}
$$

Indeed,

$$
E[< W(1), D\varphi(X_o)-X_o >] =< E[W(1)], E[D\varphi(X_o)-X_o] >= 0.
$$

For any $R > 0$, taking $f_R \in C(\mathbf{R}^d : [0,1])$ for which $f_R(x) = 1$ ($|x| \le R$) and $f_R(x) = 0$ ($|x| \ge R+1$),

$$
\begin{aligned}
& E[< X_\varepsilon(1), D\varphi(X_o)-X_o >] \\
= \;& E[< X_\varepsilon(1), D\varphi(X_o)-X_o > \{1 - f_R(X_\varepsilon(1)) f_R(D\varphi(X_o)-X_o)\}] \\
&\quad + E[< X_\varepsilon(1), D\varphi(X_\varepsilon(0)) - X_\varepsilon(0) > \\
&\qquad \times f_R(X_\varepsilon(1)) f_R(D\varphi(X_\varepsilon(0)) - X_\varepsilon(0))].
\end{aligned}
$$

$$
\begin{aligned}
& E[| < X_\varepsilon(1), D\varphi(X_o)-X_o > | \{1 - f_R(X_\varepsilon(1)) f_R(D\varphi(X_o)-X_o)\}] \\
\le \;& \sqrt{E[|D\varphi(X_o)-X_o|^2] E[|X_\varepsilon(1)|^2 : |X_\varepsilon(1)| \ge R]} \\
&\quad + \sqrt{E[|X_\varepsilon(1)|^2] E[|D\varphi(X_o)-X_o|^2 : |D\varphi(X_o)-X_o| \ge R]} \to 0
\end{aligned}
$$

as $R \to \infty$, uniformly in $\varepsilon \in [0,1]$. Since $(X_\varepsilon(0), X_\varepsilon(1))$ weakly converges to $(X_o, D\varphi(X_o))$ as $\varepsilon \to 0$ by the uniqueness of the minimizer of $V(P_0,P_1)$, one can assume, by taking a new probability space $(\tilde{\Omega}, \tilde{\mathbf{B}}, \tilde{P})$, that $(X_\varepsilon(0), X_\varepsilon(1))$ converges to $(X_o, D\varphi(X_o))$ as $\varepsilon \to 0$, $\tilde{P}$-a.s., by Skhorohod's theorem (see e.g. [12, p. 9]). Put

$$
A := \{y \in \mathbf{R}^d | \varphi(y) < \infty, \partial\varphi(y) = \{D\varphi(y)\}\}.
$$

Then $X_o \in A$ a.s. from (A.1) and $\cap_{r>0} \partial\varphi(U_r(x)) = \{D\varphi(x)\}$ for any $x \in A$ (see [25, p. 54]), from which the following holds:

$$
\begin{aligned}
& E[< X_\varepsilon(1), D\varphi(X_\varepsilon(0)) - X_\varepsilon(0) > f_R(X_\varepsilon(1)) f_R(D\varphi(X_\varepsilon(0)) - X_\varepsilon(0))] \\
\to \ & \tilde{E}[< D\varphi(X_o), D\varphi(X_o) - X_o > f_R(D\varphi(X_o)) f_R(D\varphi(X_o) - X_o) : X_o \in A] \\
& \qquad (\text{as } \varepsilon \to 0) \\
\to \ & E[< D\varphi(X_o), D\varphi(X_o) - X_o >] \quad (\text{as } R \to \infty).
\end{aligned}
$$

(3.28)-(3.30) imply (2.5)-(2.6).

Q. E. D.

We give technical lemmas and then prove Proposition 2.1.

Lemma 3.3 *(see [17, Lemma 2.5]). Suppose that (A.2) holds. Then for any* $\varepsilon > 0$,

$$
V_\varepsilon(P_0, P_1) = 2\varepsilon E\Big[\log\Big\{\frac{\overline{h}_\varepsilon(1, \overline{X}_\varepsilon(1))}{\overline{h}_\varepsilon(0, \overline{X}_\varepsilon(0))}\Big\}\Big] \tag{3.31}
$$

(see above Proposition 2.1 for notation). $V_\varepsilon(P_0, P_1)$ *is also the infimum of*

$$
\int_0^1 \int_{\mathbf{R}^d} |b(t,x)|^2 q(t,x) dt dx \tag{3.32}
$$

over all (b, q) *for which*

$$
q(t,x) \geq 0 \quad dx - a.e., \quad \int_{\mathbf{R}^d} q(t,x) dx = 1 \quad \text{for all } t \in [0,1], \tag{3.33}
$$

$$
q(0,x)dx = P_0(dx), \quad q(1,x)dx = P_1(dx), \tag{3.34}
$$

and for which the following holds: for any $f \in C_o^\infty(\mathbf{R}^d)$ *and any* $t \in [0,1]$,

$$
\begin{aligned}
& \int_{\mathbf{R}^d} f(x)(q(t,x) - q(0,x)) dx \qquad (3.35) \\
= \ & \int_0^t ds \int_{\mathbf{R}^d} \Big(\frac{\varepsilon}{2} \triangle f(x) + < b(t,x), Df(x) > \Big) q(s,x) dx,
\end{aligned}
$$

where $\triangle := \sum_{i=1}^d \partial^2 / \partial x_i^2$.

Remark 3.3 *Suppose that (A.1) and (A.2) hold and that* $supp(P_0) \cup supp(P_1)$ *is bounded. Then it is known that* $\tilde{V}(P_0, P_1)$ *is the infimum of (3.32) over all* (b, q) *for which (3.33)-(3.35) hold for* $\varepsilon = 0$ *and for which* $\cup_{0 \le t \le 1} supp(q(t, \cdot))$ *is bounded (see [5] or [25, p. 239]).*

Lemma 3.4 *Suppose that (A.0), (A.2) and (A.3) hold. Then for any* $\varepsilon > 0$, $V_\varepsilon(P_0, P_1)$ *is finite. In particular,* $V_1(P_{1,1}, P_1)$ *is finite.*

Proof. Replace μ_ε by $\overline{\mu}_\varepsilon$ (see above Proposition 2.1 for notation) in (3.1) and denote by $\overline{H}_{m,\varepsilon}$ a function obtained from (3.1). Then, from (3.2), (3.4), (3.8), (3.31) and (A.2),

$$\begin{aligned} V_\varepsilon(P_0, P_1) &= E[|\overline{X}_\varepsilon(0)|^2 + |\overline{X}_\varepsilon(1)|^2 - 2\overline{H}_{\infty,\varepsilon}(\overline{X}_\varepsilon(0), \overline{X}_\varepsilon(1))] \\ &\quad +2\varepsilon \int_{\mathbf{R}^d} (\log p_1(x)) P_1(dx) + 2\varepsilon \log(2\pi\varepsilon)^{d/2}. \end{aligned} \tag{3.36}$$

From (3.1), (3.7), (3.13) and (A.0), for sufficiently large $m \ge 1$,

$$E[\overline{H}_{\infty,\varepsilon}(\overline{X}_\varepsilon(0), \overline{X}_\varepsilon(1))] \ge E[\overline{H}_{m,\varepsilon}(\overline{X}_\varepsilon(0), \overline{X}_\varepsilon(1))] > -\infty. \tag{3.37}$$

This together with (A.3) completes the proof.

Q. E. D.

(Proof of Proposition 2.1). The proof of (i) and (ii) is almost the same as that of Corollary 2.2. Most parts of the proof of (iii) is almost the same as that of Theorem 2.2. In fact, from Lemma 3.4, the only thing we have to prove is the following:

$$\limsup_{\varepsilon \to 0} V_\varepsilon(P_0, P_1) \le V(P_0, P_1). \tag{3.38}$$

Take ψ for which $P_0 \psi^{-1} = P_1$, which is possible from Corollary 2.1. For $r \in (0, 1/2)$, solve Schrödinger's functional equation:

$$\begin{aligned} P_{1,\varepsilon(1-r)}(dx) &= \left(\int_{\mathbf{R}^d} g_{\varepsilon r}(x - y) \nu_{1,r,\varepsilon}(dy) \right) \nu_{0,r,\varepsilon}(dx), \\ P_1(dy) &= \left(\int_{\mathbf{R}^d} g_{\varepsilon r}(x - y) \nu_{0,r,\varepsilon}(dx) \right) \nu_{1,r,\varepsilon}(dy). \end{aligned} \tag{3.39}$$

For $t \in [0, 1-r]$, put

$$X_{r,\varepsilon}(t) := X_o + t\frac{\psi(X_o) - X_o}{1-r} + \sqrt{\varepsilon}W(t), \tag{3.40}$$

and solve the following: for $t \in [1-r, 1]$

$$X_{r,\varepsilon}(t) = X_{r,\varepsilon}(1-r) + \int_{1-r}^{t} b_{r,\varepsilon}(s, X_{r,\varepsilon}(s))ds + \sqrt{\varepsilon}(W(t) - W(1-r)), \tag{3.41}$$

where

$$b_{r,\varepsilon}(s,x) := \varepsilon D_x \log\Big(\int_{\mathbf{R}^d} g_{\varepsilon(1-s)}(x-y)\nu_{1,r,\varepsilon}(dy)\Big).$$

Then, from Lemma 3.3,

$$V_\varepsilon(P_0, P_1) \le \frac{E[|\psi(X_o) - X_o|^2]}{1-r} + E[\int_{1-r}^{1} |b_{r,\varepsilon}(s, X_{r,\varepsilon}(s))|^2 ds] \tag{3.42}$$

since $X_{r,\varepsilon}(0) = X_o$ and $PX_{r,\varepsilon}(1)^{-1} = P_1$.

We prove the following to complete the proof: for any $r \in (0, 1/2)$,

$$\lim_{\varepsilon \to 0} E[\int_{1-r}^{1} |b_{r,\varepsilon}(s, X_{r,\varepsilon}(s))|^2 ds] = 0. \tag{3.43}$$

Put

$$p_{r,\varepsilon}(t,x) := \begin{cases} \int_{\mathbf{R}^d} g_{\frac{\varepsilon(1-r)(1-t)}{r}}(x-y)P_1(dy) & (1-r \le t < 1), \\ p_1(x) & (t=1). \end{cases} \tag{3.44}$$

Then

$$p_{r,\varepsilon}(1-r, x)dx = P_{1,\varepsilon(1-r)}(dx),$$

and $p_{r,\varepsilon}(t,x)$ is a weak solution to the following: for $t \in [1-r, 1)$,

$$\frac{\partial p_{r,\varepsilon}(t,x)}{\partial t} = \frac{\varepsilon}{2}\triangle p_{r,\varepsilon}(t,x) - div\Big\{\Big(\frac{\varepsilon}{2r}\Big)\frac{D_x p_{r,\varepsilon}(t,x)}{p_{r,\varepsilon}(t,x)} p_{r,\varepsilon}(t,x)\Big\}. \tag{3.45}$$

Hence, from Lemmas 3.3 and 3.4, for $\varepsilon < 1$,

$$\begin{aligned} & E[\int_{1-r}^{1} |b_{r,\varepsilon}(s, X_{r,\varepsilon}(s))|^2 ds] \\ \le\ & \int_{1-r}^{1} dt \int_{\mathbf{R}^d} \left| \left(\frac{\varepsilon}{2r}\right) \frac{D_x p_{r,\varepsilon}(t,x)}{p_{r,\varepsilon}(t,x)} \right|^2 p_{r,\varepsilon}(t,x) dx \\ =\ & \frac{\varepsilon}{4r(1-r)} \int_{1-\varepsilon(1-r)}^{1} ds \int_{\mathbf{R}^d} \left| \frac{D_x p_{\frac{1}{2},1}(s,x)}{p_{\frac{1}{2},1}(s,x)} \right|^2 p_{\frac{1}{2},1}(s,x) dx \to 0 \quad (\text{as } \varepsilon \to 0), \end{aligned} \tag{3.46}$$

where we used the following change of variable:

$$\frac{\varepsilon(1-r)(1-t)}{r} = 1 - s,$$

and the following:

$$\begin{aligned} & \int_{1-\varepsilon(1-r)}^{1} ds \int_{\mathbf{R}^d} \left| \frac{D_x p_{\frac{1}{2},1}(s,x)}{p_{\frac{1}{2},1}(s,x)} \right|^2 p_{\frac{1}{2},1}(s,x) dx \\ \le\ & \int_{0}^{1} ds \int_{\mathbf{R}^d} \left| \frac{D_x p_{\frac{1}{2},1}(s,x)}{p_{\frac{1}{2},1}(s,x)} \right|^2 p_{\frac{1}{2},1}(s,x) dx = V_1(P_{1,1}, P_1) < \infty. \end{aligned}$$

Q. E. D.

References

[1] Ambrosio L and Pratelli A, Existence and stability results in the L^1 theory of optimal transportation, preprint.

[2] Bakelman IJ, *Convex Analysis and Nonlinear Geometric Elliptic Equations* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1994).

[3] Brenier Y, Décomposition polaire et réarrangement monotone des champs de vecteurs, C. R. Acad. Sci. Paris Sr. I Math. **305,** (1987) 805–808.

[4] Brenier Y, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. **44,** (1991) 375–417.

[5] Brenier Y and Benamou JD, A numerical method for the optimal mass transport problem and related problems, in *Monge Ampre equation: applications to geometry and optimization, Proceedings of the NSF-CBMS Conference, Deerfield Beach, FL, 1997 (Caffarelli LA and Milman M, ed.), Contemporary Mathematics **226,*** (Amer. Math. Soc., Providence, RI, 1999) 1–11.

[6] Caffarelli LA, Feldman M and McCann RJ, Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs, J. of the Amer. Math. Soc. **15,** (2001) 1–26.

[7] Doob JL, *Classical potential theory and its probabilistic counterpart* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1984).

[8] Evans LC, Partial differential equations and Monge-Kantorovich mass transfer, *Current developments in mathematics, 1997 (Cambridge, MA)* (Int. Press, Boston, MA, 1999) 65–126.

[9] Evans LC and Gangbo W, Differential equations methods for the Monge-Kantorovich mass transfer problem, Mem. Amer. Math. Soc. **137,** (1999) No. 653.

[10] Fleming WH and Soner HM, *Controlled Markov Processes and Viscosity Solutions* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1993).

[11] Föllmer H, Random fields and diffusion processes, in *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87 (Hennequin PL, ed.), Lecture Notes in Math. **1362*** (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1988) 101–203.

[12] Ikeda N and Watanabe S, *Stochastic differential equations and diffusion processes* (North-Holland/Kodansha, Amsterdam, New York, Oxford, Tokyo, 1981).

[13] Jamison B, Reciprocal processes, Z. Wahrsch. Verw. Gebiete **30,** (1974) 65–86.

[14] Jamison B, The Markov process of Schrödinger, Z. Wahrsch. Verw. Gebiete **32,** (1975) 323–331.

[15] Knott M and Smith CS, On the optimal mapping of distributions, J. of Optimization Theory and Applications **43,** (1984) 39–49.

[16] Knott M and Smith CS, Note on the optimal transportation of distributions, J. of Optimization Theory and Applications **52**, (1987) 323–329.

[17] Mikami T, Variational processes from the weak forward equation, Commun. Math. Phys. **135,** (1990) 19–40.

[18] Mikami T, Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric, Appl. Math. Optim. **42,** (2000) 203–227.

[19] Mikami T, Optimal control for absolutely continuous stochastic processes and the mass transportation problem, Elect. Comm. in Probab. **7,** (2002) 199–213.

[20] Rachev ST and Rüschendorf L, *Mass transportation problems, Vol. I: Theory* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1999).

[21] Rüschendorf L and Rachev ST, A characterization of random variables with minimum L^2-distance, J. of Multivarite Analysis **32,** (1990) 48–54.

[22] Rüschendorf L and Thomsen W, Note on the Schrödinger equation and I-projections, Statist. Probab. Lett. **17,** (1993) 369–375.

[23] Salisbury TS, An increasing diffusion, in *Seminar on Stochastic Processes 1984 (Cinlar E, Chung KL, Getoor RK, eds.)* (Birkhäuser, Boston, MA, 1986) 173–194.

[24] Trudinger NS and Wang Xu-Jia, On the Monge mass transfer problem, Calc. Var. **13,** (2001) 19–31.

[25] Villani C, *Topics in Optimal Transportation, Graduate Studies in Mathematics Vol. 58* (Amer. Math. Soc., Providence, RI, 2003).

[26] Zambrini, JC, Variational processes. in *Stochastic processes in classical and quantum systems, Ascona, 1985, (Albeverio, S., Casati, G., Merlini, D., eds.), Lecture Notes in Phys.* ***262*** (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986) 517–529.